How Much Training is Needed for Interference Coordination in Cellular Networks?

Hans H. Brunner, Mario H. Castañeda, Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Technische Universität München, 80290 Munich, Germany
e-mail: {brunner|mario.castaneda|josef.a.nossek}@tum.de

***Abstract*—Cooperative techniques for cellular networks promise very high data rates, but require additional and precise knowledge of the serving and interference channels. We show, how the pilot symbols required for achieving this information affect the possible data rates. The measurements of the channels are suffering from pilot contamination, due to the measurements in adjacent cells. On the one hand, with a too short pilot length, cooperation is not possible and the channels are learned too poorly, degrading the possible data rates. On the other hand, a too long pilot length reduces the efficiency of the system, leaving no resources for the data transmission. In addition, the channel measurements are outdated before they can be applied. With an upper bound to the sum rate of a system with interference coordination and a sub-optimal pilot allocation strategy, we discuss the pilot length trade-off.**

I. Introduction

It is discussed, whether *cooperative multi-point* (CoMP) techniques should be included in future standards. The benefits of CoMP sound very promising, but it is not proved so far that cooperation will be beneficial, if all implementation issues are taken into account. Under the term CoMP, we understand any transmission technique, which allows more than one link per user. This includes, e.g., soft handover, where a *mobile device* (MD) is connected continuously to multiple *base stations* (BSs) and is dynamically served over the strongest link. But, it also includes *network multiple input multiple output* (network MIMO), where the whole network can be seen as a single broadcast channel with distributed transmit antennas.

In this contribution, we restrict cooperation to interference coordination, where each MD is only served by one BS. But, the BSs are allowed to control the interference they produce, e.g., as done with interference alignment. Our analysis is based on an upper bound, therefore, we do not specify the cooperation any further. The existence of a backhaul network connecting all BSs is most likely, but not strictly necessary.

Usually, cooperation aims at mitigating *intercell interference* (ICI), which is the most important effect limiting the possible data rates in cellular systems [1]. State-of-the-art systems try to cope with this problem by using static resource splitting like frequency reuse and sectorization. In a system with frequency reuse, the ICI can be reduced substantially, but a BS can only use a fraction of the available bandwidth, which degrades the possible data rates dramatically.

For the *downlink* (DL) of a cellular system with *time division duplex* (TDD), we describe a limit to beneficial cooperation. The *channel state information* (CSI) required for the cooperative transmission technique defines a minimum uplink pilot length. The pilot length influences the system efficiency and, therefore, minders the possible data rates. With an upper bound to the network sum rate with interference coordination, an optimal level of cooperation and the corresponding pilot length can be found. In contrast to [2], we also take pilot contamination [3] or ICI during the channel measurements into account.

On the one hand, we discuss in Section V how to allocate pilot sequences to MDs with respect to pilot contamination for a given pilot length. And with a given pilot contamination on the other hand, we identify an optimal pilot length with an upper bound to the network sum rate with interference coordination in Section IV. An iterative algorithm alternating between these two optimizations is described in Section VI. In Section VII, Monte Carlo simulations for the described methods are discussed and limits to beneficial cooperation identified.

II. System Model

We consider a cellular network with 19 three faced sites and, therefore, 57 BSs. Each BS serves the MDs of the hexagonal shaped cell it covers. A MD in the set $\mathcal{K}$ of all MDs is specified by the tuple $(b, k) \in \mathcal{K}$, where $b \in \mathcal{B}$ identifies the BS in the set $\mathcal{B} = \{1, \ldots, 57\}$ of all BSs and $k \in \mathcal{K}_b$ the MD in the set $\mathcal{K}_b = \{1, \ldots, K\}$ of all MDs in the cell of BS b. In this paper, each BS has N antennas and serves $K = |\mathcal{K}_b|$ single antenna MDs, respectively.

The wrap-around method is used to treat all cells equally. The 57 BSs are copied, including their beamforming, and placed six times around the central cluster. Each MD only sees the 57 BSs, which are closest by Euclidean distance. In Figure 1, the cellular layout can be seen, where the central cluster is inked slightly darker than the wrap-a-round clusters. The placement and orientation of the BSs is indicated by small arrows.

A. Channel Model

The spatial channel model of the 3GPP MIMO urban macro cell with a distance of 500m between the two closest sites and a center frequency of 2 GHz is utilized [4]. The applied parameters for the simulations can be seen in Table I.

978-1-61284-074-1/11/$26.00

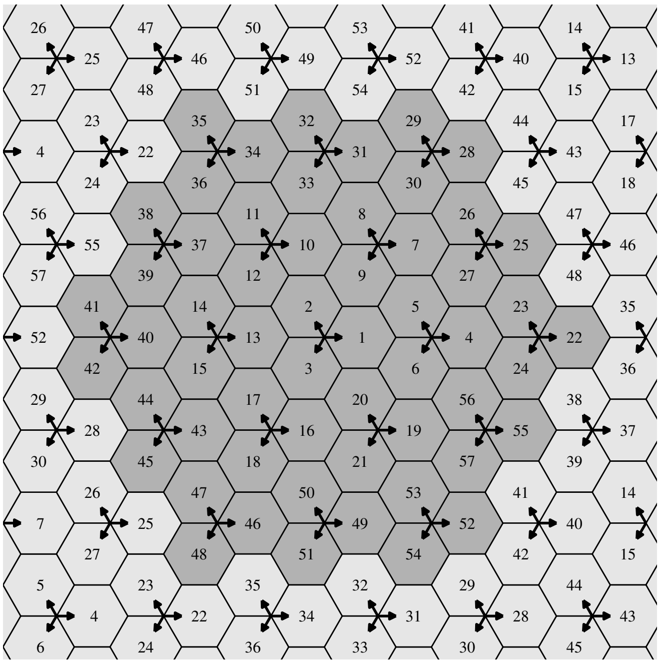

Figure 1. Cellular Cluster with Wrap-a-Round

scenario	urban macro-cell
center frequency	2GHz
sectors	$19 \cdot 3 = 57$
inter site distance	500m
users per sector	$K = \lvert\mathcal{K}_b\rvert$
min distance to site	25m
antenna configuration	$N \times 1$ MISO
antenna spacing	0.5λ
user speed	slow: $3\,\text{km/h}$ fast: $30\,\text{km/h}$
root mean square delay spread	$\sigma_{DS} = 0.5\,\mu\text{s}$
transmit power	BS: $P = 83\,\text{W}$ MD: $P_{\text{MD}} = 0.83\,\text{W}$
background interference	$\theta^2_{\text{bg}} = 1.98 \cdot 10^{-11}\,\text{W}$
thermal noise	$\sigma^2_\eta = 8.3 \cdot 10^{-14}\,\text{W}$

Table I
SIMULATION PARAMETERS

The vectors $\boldsymbol{h}_{\hat{b},b,k} \in \mathbb{C}^N$ contain the channel coefficients between the antennas of BS $\hat{b}$ and MD (b,k). With $(\bullet)^{\text{T}}$ and $(\bullet)^{\text{H}}$ we denote the transposition and the complex conjugate transposition, respectively. The achievable, normalized rate with *dirty paper coding* (DPC) of MD (b,k) can be expressed as

$$r_{b,k} = \log_2\left(1 + \frac{|\boldsymbol{h}^{\text{T}}_{b,b,k}\boldsymbol{p}_{b,k}|^2}{\sigma^2_{b,k} + \sum_{\hat{k}<k}|\boldsymbol{h}^{\text{T}}_{b,b,k}\boldsymbol{p}_{b,\hat{k}}|^2 + \theta^2_{b,k}}\right), \quad (1)$$

$$\theta^2_{b,k} = \sum_{\hat{b}\in\mathcal{B}\setminus b} \boldsymbol{h}^{\text{H}}_{\hat{b},b,k}\boldsymbol{Q}_{\hat{b}}\boldsymbol{h}_{\hat{b},b,k} + \theta^2_{\text{bg}}, \quad (2)$$

where $\boldsymbol{p}_{b,k} \in \mathbb{C}^N$ is the beamforming vector for MD (b,k) and $\boldsymbol{Q}_b = \sum_k \boldsymbol{p}_{b,k}\boldsymbol{p}^{\text{H}}_{b,k} \in \mathbb{C}^{N\times N}$ is the sum transmit covariance matrix of BS b. $\sigma^2_{b,k}$ is the variance of the noise, $\sum_{\hat{k}<k}|\boldsymbol{h}^{\text{T}}_{b,b,k}\boldsymbol{p}_{b,\hat{k}}|^2$ is the variance of the intracell interference with DPC, and $\theta^2_{b,k}$ is the variance of the received intercell interference. The BSs further away than the closest 57 BSs are modeled by a Gaussian background interference θ^2_{bg} for a given signal variance per transmit antenna. All BSs have to satisfy the transmit power constraint $\text{tr}(\boldsymbol{Q}_b) \leq P$, $\forall b$.

B. TDD Signaling

In [5], it was shown that we can use the pilot length of an idealized TDD system as a lower bound for a frequency division duplex system. In TDD systems, the reciprocity of the propagation channels is exploited. The signal processing in the downlink is based on the channel estimation in the *uplink* (UL). In Figure 2, the signaling scheme can be seen. The MDs send a pilot sequence of length T_{pilot} in the UL, which can be used by the BSs to measure the channels. Based on this informaiton, the BSs optimize their beamforming, which they utilize to transmit a second pilot of length T_{2nd} [6]. With this second pilot, the MDs can measure their effecftive *signal to interference plus noise ratio* (SINR), which they feed back to the BSs during $T_{\text{sinr fb}}$. The BSs do not change their beamforming based on the updated SINR values, but can serve the MDs with interference aware rates during the data transmission of length T_{data}.

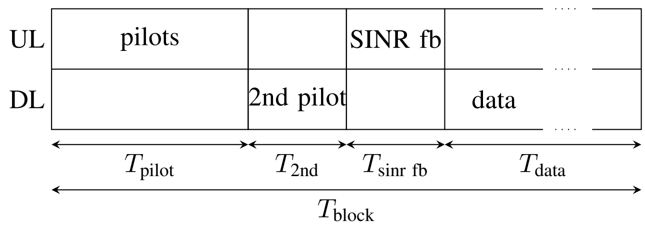

Figure 2. TDD Signaling

We take the pilot length T_{pilot} as the dominant overhead and neglect T_{2nd}, $T_{\text{sinr fb}}$ and any other overhead. The block length $T_{\text{block}} = T_{\text{pilot}} + T_{\text{data}}$ is defined by the recurrence of channel measurements. The efficiency of this signaling scheme is $T_{\text{data}}/T_{\text{block}}$. We assume, that piloting and the selection of the beamforming vectors are synchronized at all BSs.

III. PILOT CONTAMINATION

The MDs are split into equally sized disjoint subsets $\mathcal{S}_s$. The MDs within a subset use pilot sequences, which are orthogonal to each other. But, these pilot sequences are reused in all other subsets. Therefore, the measurement of a channel is always disturbed by the transmission of MDs using the same pilot sequence. This interference during the pilot phase is also called pilot contamination. Note, that the pilot length has to be at least as large as the number of users in a subset $T_{\text{pilot}} \geq |\mathcal{S}_s|$, $\forall s$.

The K MDs within a cell have to use orthogonal pilot sequences, in order to have the associated BSs obtain meaningful estimates. A pilot length of $T_{\text{pilot}} = K + L$ allows up to $K + L$ orthogonal pilot sequences. Each BS can measure the channels to its own K MDs and L interference channels, additionally. Using all possible pilot sequences is always beneficial, as it allows to learn L interference channels at each BS, respectively, without additional costs and minimizes the pilot contamination. We set $T_{\text{pilot}} = |\mathcal{S}_s| = K + L, \forall s$.

With a common MD transmit power P_{MD} and the disjoint sets $\mathcal{P}_p$ of all MDs sharing the same pilot sequence with index p, such a pilot allocation will result for an MD in the set $\mathcal{P}_p$ in the contamination at BS $\hat{b}$ for MD (b, k)

$$\boldsymbol{\Theta}_{\text{pc},\hat{b},b,k} = P_{\text{MD}} T_{\text{pilot}} \sum_{(\tilde{b},\tilde{k}) \in \mathcal{P}_p \setminus (b,k)} \boldsymbol{R}_{\boldsymbol{h},\hat{b},\tilde{b},\tilde{k}}, \tag{3}$$

$$= P_{\text{MD}} T_{\text{pilot}} \left(\boldsymbol{R}_{\boldsymbol{h},\hat{b},p} - \boldsymbol{R}_{\boldsymbol{h},\hat{b},b,k} \right), \tag{4}$$

where $\boldsymbol{R}_{\boldsymbol{h},\hat{b},b,k} = \mathrm{E}\left[\boldsymbol{h}_{\hat{b},b,k} \boldsymbol{h}_{\hat{b},b,k}^{\mathrm{H}} \right]$ is the covariance matrix of channel $\boldsymbol{h}_{\hat{b},b,k}$ and $\boldsymbol{R}_{\boldsymbol{h},\hat{b},p} = \sum_{(\tilde{b},\tilde{k}) \in \mathcal{P}_p} \boldsymbol{R}_{\boldsymbol{h},\hat{b},\tilde{b},\tilde{k}}$ is the sum of the covariance matrices of all channels from MDs using the pilot sequence with index p to the BS $\hat{b}$. As described in [7] and [8] for the point-to-point and the broadcast channel, respectively, the measurement error due to noise and interference during the piloting can be modeled as Gaussian noise for the data transmission. With an MMSE estimator, the error covariance matrix due to contamination is

$$\boldsymbol{R}_{\text{pc},\hat{b},b,k} = \left(\boldsymbol{R}_{\boldsymbol{h},\hat{b},b,k}^{-1} + P_{\text{MD}} T_{\text{pilot}} \left(\sigma_\eta^2 \mathbf{I} + \boldsymbol{\Theta}_{\text{pc},\hat{b},b,k} \right)^{-1} \right)^{-1} \tag{5}$$

$$= \boldsymbol{R}_{\boldsymbol{h},\hat{b},b,k} + \boldsymbol{R}_{\boldsymbol{h},\hat{b},b,k} \left(\frac{\sigma_\eta^2}{P_{\text{MD}} T_{\text{pilot}}} \mathbf{I} + \boldsymbol{R}_{\boldsymbol{h},\hat{b},p} \right)^{-1} \boldsymbol{R}_{\boldsymbol{h},\hat{b},b,k}, \tag{6}$$

where σ_η^2 is the thermal noise.

IV. Upper Bound to Interference Coordination

With L measured interference channels per BS, an upper bound to interference coordination can be given, as described in [2]. The intercell interference (2) can be split into

$$\theta_{b,k}^2 = \sum_{\hat{b} \in \mathcal{C}_{b,k} \setminus b} \underbrace{\hat{\boldsymbol{h}}_{\hat{b},b,k}^{\mathrm{H}}}_{\text{known}} \boldsymbol{Q}_{\hat{b}} \hat{\boldsymbol{h}}_{\hat{b},b,k} + \sum_{\hat{b} \in \mathcal{B} \setminus \mathcal{C}_{b,k}} \underbrace{\boldsymbol{h}_{\hat{b},b,k}^{\mathrm{H}}}_{\text{unknown}} \boldsymbol{Q}_{\hat{b}} \boldsymbol{h}_{\hat{b},b,k} + \underbrace{\theta_{\text{bg}}^2}_{\text{known}} \tag{7}$$

$$= \sum_{\hat{b} \in \mathcal{C}_{b,k} \setminus b} \underbrace{\hat{\boldsymbol{h}}_{\hat{b},b,k}^{\mathrm{H}}}_{\text{known}} \boldsymbol{Q}_{\hat{b}} \hat{\boldsymbol{h}}_{\hat{b},b,k} + \underbrace{\theta_{\text{no,b,k}}^2}_{\text{known}}. \tag{8}$$

$\hat{\boldsymbol{h}}_{\hat{b},b,k}$ is the estimated channel between BS $\hat{b}$ and MD (b, k). The set $\mathcal{C}_{b,k}$ contains all BSs, which know the channel to the MD (b, k). Therefore, $\sum_{\hat{b} \in \mathcal{C}_{b,k} \setminus b} \hat{\boldsymbol{h}}_{\hat{b},b,k}^{\mathrm{H}} \boldsymbol{Q}_{\hat{b}} \hat{\boldsymbol{h}}_{\hat{b},b,k}$ is the sum of the interference over all measured and, therefore, known interference channels. The rest of the terms in (7) is the sum interference over the unknown interference channels including the background interference. We assume to know the variance $\theta_{\text{no,b,k}}^2$ from SINR measurements, but we cannot optimize over the precoding vectors transmitting over these unknown interference channels.

We include *channel state information* (CSI) outdating with the common normalized mean outdating error variance σ_{od}^2 associated with a given block length [2]. The measurement error covariance matrix combines to $\boldsymbol{R}_{e,\hat{b},b,k} = \sigma_{\text{od}}^2 \boldsymbol{R}_{\boldsymbol{h},\hat{b},b,k} + \boldsymbol{R}_{\text{pc},\hat{b},b,k}$. This still neglects the fact that the TDD UL measurements for the DL will always contain a calibration error because the reciprocity is only an idealized assumption. For the noise term of a MD, we add the measurement errors for the serving and the nulled interference channels to the thermal noise noise term

$$\sigma_{b,k}^2 = \sigma_\eta^2 + \mathrm{tr}(\boldsymbol{R}_{\text{e},b,b,k} \boldsymbol{Q}_b) + \sum_{\hat{b} \in \mathcal{C}_{b,k} \setminus b} \mathrm{tr}(\boldsymbol{R}_{\text{e},\hat{b},b,k} \boldsymbol{Q}_{\hat{b}}). \tag{9}$$

With either the assumption that the measurement error covariances or the transmit covariances in the noise term are scaled identity matrices, this noise term does not depend on the used precoding vectors.

The jointly optimized sum rate R_{coop} is always smaller than the upper bound R_{upper}, where all measured interference channels are set to zero:

$$R_{\text{coop}}(L) \le R_{\text{upper}}(L) = \frac{T_{\text{data}}}{T_{\text{block}}} \max_{\{\boldsymbol{p}_{b,k} | \forall (b,k) \in \mathcal{K}\}} \sum_{(b,k) \in \mathcal{K}} \hat{r}_{b,k}, \quad \text{s.t. } \mathrm{tr}(\boldsymbol{Q}_b) \le P \ \forall b, \tag{10}$$

$$\hat{r}_{b,k} = r_{b,k}|_{\hat{\boldsymbol{h}}_{\hat{b},b,k} = \mathbf{0} \ \forall \hat{b} \in \mathcal{C}_{b,k} \setminus b} = \log_2 \left(1 + \frac{|\hat{\boldsymbol{h}}_{b,b,k}^{\mathrm{T}} \boldsymbol{p}_{b,k}|^2}{\sigma_{b,k}^2 + \sum_{\hat{k} < k} |\hat{\boldsymbol{h}}_{b,b,k}^{\mathrm{T}} \boldsymbol{p}_{b,\hat{k}}|^2 + \theta_{\text{no},b,k}^2} \right). \tag{11}$$

Problem (10) is convex and can be solved distributed at all BSs independently with iterative water-filling [9]. With the given CSI, R_{upper} is a loose upper bound, which is always maximized with the maximum transmit power. Finding precoding vectors resulting in a higher sum rate would require additional measurements. The upper bound is not achievable, because the cost of nulling the L interference channels per BS is neglected.

V. Pilot Allocation

The selection of L is the central focus of this paper. For a given L, the $K + L$ pilot sequences have to be allocated to the MDs in such a way, that a BS can at least measure the channels to the MDs it serves. Additionally, L interference channels can be estimated at each BS, respectively.

In the following, we will discuss different strategies for allocating the pilot sequences to the MDs with respect to the pilot contamination.

A. Random Allocation

For the random pilot sequence allocation strategy, we assign the pilot sequences to the MDs based on their index. The BSs are randomly labeled with an index b and the MDs associated to BS b are also randomly labeled with the index k. The set of all MDs, which use the pilot sequence with index p is

$$\mathcal{P}_p^{\text{rand}} = \{(b,k) : (b,k) \in \mathcal{K}, p = \text{mod}((b-1)K + k, T_{\text{pilot}})\}, \tag{12}$$

where $\text{mod}(a,b) = a - b\left\lfloor \frac{a}{b} \right\rfloor + 1$ is a slightly modified division algorithm. This assures that all MDs within one cell use a different pilot sequence, as long as $T_{\text{pilot}} \geq K$. Therefore, each BS can measure the channels of the associated MDs utilizing K of the pilot sequences. For each of the remaining L pilot sequences, each BS can measure the strongest of the interference channels linking to MDs with the specific pilot sequence, respectively.

B. Strongest Interferer Allocation

The upper bound in Section IV reaches the highest rates, if the L known interference channels at each BS are the channels over which the BSs would cause the strongest decrease in rate, respectively. Finding these optimal interference channels is a problem, which is very hard to tackle. Therefore, we approach the optimum by suboptimally choosing the L interference channels at each BS, which have the largest Euclidean norm, respectively.

To identify the channels a BS is supposed to measure, we create for each BS the set $\mathcal{L}_b$, which contains the K MDs associated to the BS b and the L MDs with the strongest interference channels linked to BS b, respectively. The MDs in a set $\mathcal{L}_b$ should use pilot sequences, which are orthogonal to each other. Otherwise, the BSs could not get meaningful estimates of the channels. We want to find the sets $\mathcal{P}_p$, which minimize the trace of the the sum of all pilot contaminations, whereas each pilot sequence has to appear in each set $\mathcal{L}_b$

$$\left\{\mathcal{P}_1^{\text{opt}}, \ldots, \mathcal{P}_{T_{\text{pilot}}}^{\text{opt}}\right\} = \underset{\{\mathcal{P}_1,\ldots,\mathcal{P}_{T_{\text{pilot}}}\}}{\text{argmin}} \sum_{\hat{b}\in\mathcal{B}} \sum_{(b,k)\in\mathcal{L}_{\hat{b}}} \text{tr}\left(\boldsymbol{\Theta}_{\text{pc},\hat{b},b,k}\right)$$
$$\text{s.t. } \mathcal{P}_p \cap \mathcal{L}_b \neq \emptyset \; \forall p, b. \tag{13}$$

The different sets $\mathcal{L}_b$ for the different BSs are overlapping. Therefore, it is not always feasible to allocate the pilot sequences to the MDs in such a way, that in every set $\mathcal{L}_b$ the MDs use all different pilot sequences. To get an always feasible optimization, we relax the minimization (13) and allow different sets $\mathcal{P}_p^{\hat{b},b,k}$ of MDs using the same pilot sequence for each channel estimation, respectively. Therefore, there are no unique sets $\mathcal{P}_p$ any more and they have to be exchanged in equation (3) with $\mathcal{P}_p^{\hat{b},b,k}$ for each channel measurement.

$$\mathcal{P}_p^{\hat{b},b,k} = \underset{\mathcal{P}_p}{\text{argmin}}\, \text{tr}\left(\boldsymbol{\Theta}_{\text{pc},\hat{b},b,k}\right)$$
$$\text{s.t. } \mathcal{P}_p \cap \mathcal{L}_{\tilde{b}} \neq \emptyset \; \forall \tilde{b}. \tag{14}$$

For each channel measurement, the optimization will generate a different pilot sequence allocation. This is clearly not implementable, as the different pilot allocations may conflict with each other. An MD may be assigned to multiple pilot sequences, which is not possible.

```
for b̂ ∈ B do
  for (b,k) ∈ L_b̂ do
    P_p^{b̂,b,k} ← {(b,k)}
    for b̃ ∈ B \ b̂ do
      if P_p^{b̂,b,k} ∩ L_b̃ = ∅ then
        P_p^{b̂,b,k} ← P_p^{b̂,b,k} ∪ argmin_{(b̄,k̄)∈L_b̃ \ L_b̂} tr(R_{h,b̂,b̄,k̄})
      end if
    end for
  end for
end for
```

Table II
STRONGEST INTERFERER ALLOCATION

We use an successive allocation algorithm to solve problem (14) suboptimally. The pseudocode of the algorithm can be seen in Table II. The algorithm is initialized by setting $\mathcal{P}_p^{\hat{b},b,k} = \{(b,k)\}$. In a random order, we visit every set $\mathcal{L}_{\tilde{b}}$ and check $\mathcal{P}_p^{\hat{b},b,k} \cap \mathcal{L}_{\tilde{b}}$. If the sets have a non empty intersection, i.e., an MD in the set $\mathcal{L}_{\tilde{b}}$ is already assigned to the same pilot as MD (b,k), this set is skipped. Otherwise, all MDs in the set $\mathcal{L}_{\hat{b}}$ are removed from the set $\mathcal{L}_{\tilde{b}}$ because they have to use a different pilot sequence than MD (b,k). In the remaining set $\mathcal{L}_{\tilde{b}} \setminus \mathcal{L}_{\hat{b}}$, we pick the MD with the weakest channel linked to BS $\hat{b}$. If we would remove all MDs from previously visited sets, the constraint $\mathcal{P}_p^{\hat{b},b,k} \cap \mathcal{L}_{\tilde{b}} \neq \emptyset \; \forall \tilde{b}$ might not be fulfilled.

On the one hand, because the sets $\mathcal{L}_b$ overlap, it is not always optimal to chose the MDs with the weakest channel in the interfering sets. It might be better to chose a MD with a stronger channel, which appears in multiple sets and, therefore, reduces the number of MDs in the set $\mathcal{P}_p^{\hat{b},b,k}$. On the other hand, an implementable solution would force us to allocate the pilot sequences to the MDs without any conflicts and, therefore, reduce the degrees of freedom. We take this approach to get an insight on the benefit of an optimized pilot allocation, where each BS can measure the MDs of its cell and the L strongest interferers.

VI. Iterative Pilot Length Selection

The optimal length of the pilot sequences for a specific channel realization can be found with an iterative algorithm. For a pilot length L^ℓ and a pilot allocation strategy, we can find the measurement error covariance matrix $\boldsymbol{R}^\ell_{\text{pc},\hat{b},b,k}$ from (6). By modeling the measurement error as Gaussian noise and with a given block length, we chose the pilot length $L^{\ell+1} = \underset{L}{\text{argmax}}\, R_{\text{upper}}(L)$, which maximizes the upper bound (10). With ℓ as iterator, we alternate between these two optimizations until convergence is approached and an optimal pilot length ist found.

VII. Simulations

For different MD speeds, the normalized average cell sum rate is plotted over the block length in the following. The results are obtained with Monte Carlo simulations with 100,000 realizations, respectively. Every BS has $N = 4$

transmit antennas and transmits with $P = 83\,\mathrm{W}$. In every cell, $K = 10$ MDs are placed uniformly distributed and suffer from a thermal noise variance of $\sigma_\eta^2 = 8.3 \cdot 10^{-14}\,\mathrm{W}$ and a background interference of $\theta_{\text{bg}}^2 = 1.98 \cdot 10^{-11}$ W at their receive antenna, respectively. The MDs transmit the UL pilots with $P_{\text{MD}} = 0.83\,\mathrm{W}$. The root mean square delay spread is set to $\sigma_{DS} = 0.5\,\mu\mathrm{s}$ and the maximum Doppler frequency $f_{\text{D}} = f_{\text{C}} \frac{v}{c}$, with the center frequency f_{C} and the speed of light c, directly depends on the MD speed v. The low mobility scenarios are computed with a common MD device speed of $v = 3\,\mathrm{km/h}$ and the high mobility scenarios with $v = 30\,\mathrm{km/h}$. To solve the optimization in (10), the worst case Gaussian noise assumption for $\theta_{\text{no},b,k}^2$ and $\sigma_{b,k}^2$ are assumed.

For different choices of L and both pilot sequence allocation strategies, Figure 3 shows the locally optimal sum rate per cell (local), where no channels are set to zero. This is not the upper bound to cooperation, the intercell interference is completely regarded as noise. The different selections of L, which determine the pilot sequence length, are only used for reducing the pilot contamination. All curves ascend in the beginning for longer block lengths, because the efficiency of the system improves. At some point, each curve starts to descend, because the outdating of the channel degrades the possible rates. With the strongest interferer allocation (str), the selection $L = 0$ and a pilot sequence length of K is sufficient. The probability of a strong interferer is already so low, that a larger L does not improve the sum rate. With the random allocation strategy (rnd), the sum rate is always smaller. A pilot sequence length slightly larger than K results in the peak sum rate. But, many choices of L, including $L = 0$, will lead to very similar rates.

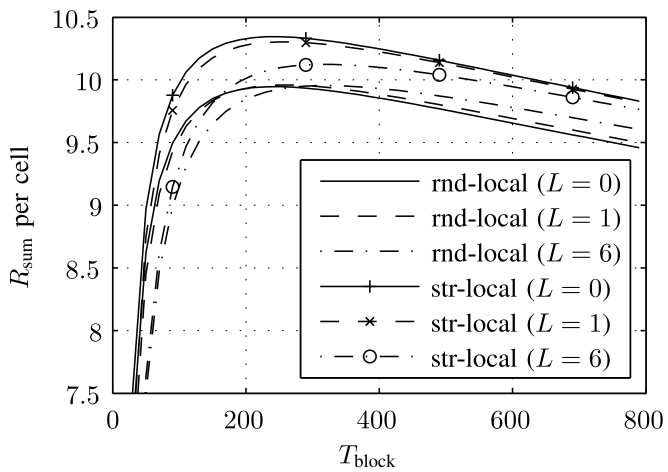

Figure 3. Low Mobility: Pilot Allocation without Cooperation

Figure 4 shows the upper bound R_{sum} per cell (coop) for the random allocation strategy and Figure 5 for the strongest interferer allocation. The selection of L does not only determine the pilot contamination, but also the degree of cooperation. L stands for the number of interference channels, which are set to zero per BS. It can be seen, that already $L = 5$ shows significant improvements compared to no cooperation ($L = 0$). For the random allocation L between 20 and 30 gives approximately 1.5 times larger rates than without cooperation. With the strongest interferer allocation L should be between 16 and 26. The strongest interferer strategy performs always better than the random allocation strategy and requires a smaller L for reaching the peak. Note, that we do not take the costs for mitigating the interference over the measured L interference channels into account. Therefore, this upper bound cannot be reached.

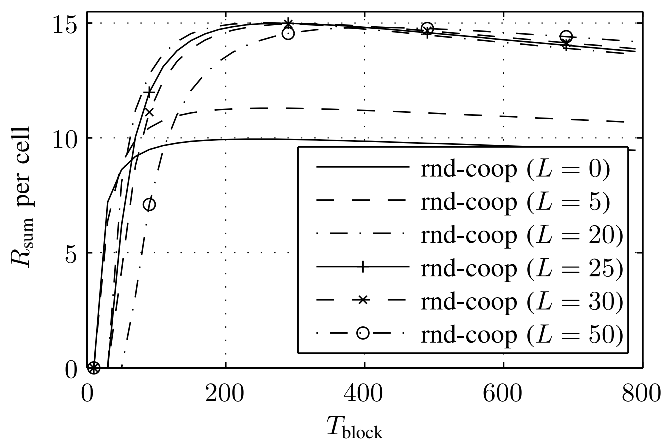

Figure 4. Low Mobility: Random Pilot Allocation with Cooperation

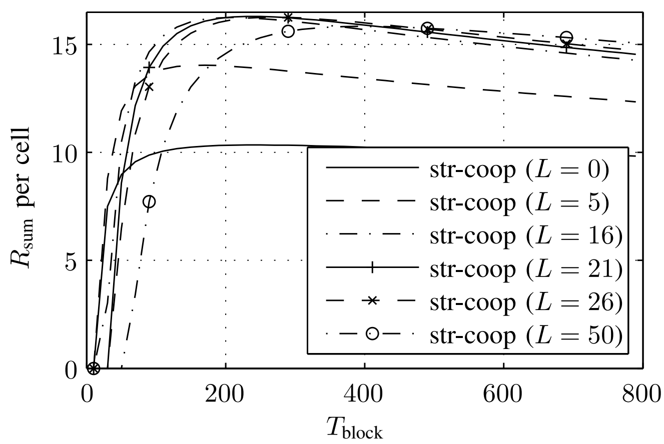

Figure 5. Low Mobility: Strongest Interferer Allocation with Cooperation

The high mobility scenario can be seen in Figure 6. The different allocation strategies with and without cooperation are compared with their optimal L. The possible improvements through cooperation are much smaller than in the low mobility scenario. L should be selected much smaller than in the low mobility scenario. The influence of the pilot allocation strategy is not very strong.

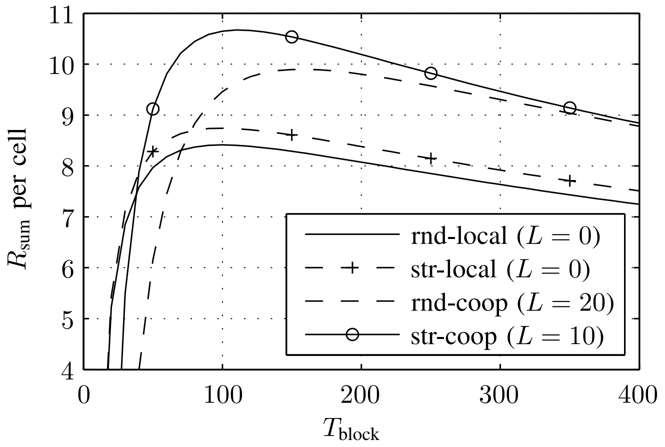

Figure 6. High Mobility: Comparison of the Peaks with Optimal L

VIII. Conclusion

Based on pilot contamination during the channel measurements, different pilot allocation strategies, and CSI outdating, we presented an upper bound to beneficial cooperation. We directly connected the degree of cooperation and the pilot length. For each piloting symbol, each BS can measure one channel, respectively. An increase of the pilot length above the discussed upper bound, would allow a higher level of cooperation, as more channels could be estimated. But, beyond the boundary the decrease in efficiency outweighs the benefits of cooperation. We could show, that even with an idealized and unachievable pilot allocation strategy the rates cannot be improved significantly.

References

[1] M. T. Ivrlač and J. A. Nossek, "Intercell-Interference in the Gaussian MISO Broadcast Channel," in *IEEE Global Telecommunications Conference*, 2007.

[2] H.H. Brunner, M. Ivrlac, and J.A. Nossek, "Upper Bound to Interference Coordination with Channel State Information Outdating," Submitted to European Wireless, 2011.

[3] J. Jose, A. Ashikhmin, T.L. Marzetta, and S. Vishwanath, "Pilot Contamination Problem in Multi-Cell TDD Systems," in *ISIT 2009, IEEE International Symposium on Information Theory*, 2009, pp. 2184–2188.

[4] "Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations," Tech. Rep. 25.996 V9.0.0, 3rd Generation Partnership Project, Technical Specification Group Radio Access Network, Dec. 2009.

[5] T.L. Marzetta and B.M. Hochwald, "Fast Transfer of Channel State Information in Wireless Systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1268–1278, Apr. 2006.

[6] H.H. Brunner and J.A. Nossek, "Mitigation of Intercell Interference without Base Station Cooperation," in *WSA 2010, International ITG Workshop on Smart Antennas*, Feb. 2010, pp. 1–7.

[7] B. Hassibi and B.M. Hochwald, "How Much Training is Needed in Multiple-Antenna Wireless Links?," *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, 2003.

[8] M. Kobayashi, G. Caire, and N. Jindal, "How Much Training and Feedback are Needed in MIMO Broadcast Channels?," in *ISIT 2008, IEEE International Symposium on Information Theory*, 2008, pp. 2663–2667.

[9] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum Power Iterative Water-Filling for Multi-Antenna Gaussian Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.